OIPE

PAGE: 1

RAW SEQUENCE LISTING
PATENT APPLICATION   *US/09/096,648*

DATE: 06/23/98
TIME: 07:42:03

INPUT SET: S26847.raw

> This Raw Listing contains the General
> Information Section and up to the first 5 pages.

```
 1                                  SEQUENCE LISTING          ENTERED
 2
 3    (1)     General Information:
 4
 5            (i) APPLICANT: Hadlaczky, Gyula
 6                           Szalay, Aladar
 7
 8            (ii) TITLE OF INVENTION: ARTIFICIAL CHROMOSOMES, USES THEREOF AND
 9    METHODS PREPARING ARTIFICIAL CHROMOSOMES
10
11            (iii) NUMBER OF SEQUENCES: 12
12
13            (iv) CORRESPONDENCE ADDRESS:
14               (A) ADDRESSEE: Brown, Martin, Haller & McClain
15               (B) STREET: 1660 Union Street
16               (C) CITY: San Diego
17               (D) STATE: CA
18               (E) COUNTRY: USA
19               (F) ZIP: 92101-2926
20
21            (v) COMPUTER READABLE FORM:
22               (A) MEDIUM TYPE: Diskette
23               (B) COMPUTER: IBM Compatible
24               (C) OPERATING SYSTEM: DOS
25               (D) SOFTWARE: FastSEQ Version 1.5
26
27            (vi) CURRENT APPLICATION DATA:
28               (A) APPLICATION NUMBER:US/09/096,648
29               (B) FILING DATE:
30               (C) CLASSIFICATION:
31
32            (vii) PRIOR APPLICATION DATA:
33               (A) APPLICATION NUMBER: 08/629,822
34               (B) FILING DATE: 10-APR-1996
35
36            (viii) ATTORNEY/AGENT INFORMATION:
37               (A) NAME: Seidman, Stephanie L
38               (B) REGISTRATION NUMBER: 33,779
39               (C) REFERENCE/DOCKET NUMBER: 6869-402A
40
41            (ix) TELECOMMUNICATION INFORMATION:
42               (A) TELEPHONE: 619-238-0999
43               (B) TELEFAX: 619-238-0062
44               (C) TELEX:
45
46
```

```
47            (2) INFORMATION FOR SEQ ID NO:1:
48
49        (i) SEQUENCE CHARACTERISTICS:
50          (A) LENGTH: 1293 base pairs
51          (B) TYPE: nucleic acid
52          (C) STRANDEDNESS: single
53          (D) TOPOLOGY: linear
54
55        (ii) MOLECULE TYPE: Genomic DNA
56        (iii) HYPOTHETICAL: NO
57        (iv) ANTI-SENSE: NO
58        (v) FRAGMENT TYPE:
59        (vi) ORIGINAL SOURCE:
60        (ix) FEATURE:
61
62        (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:
63
64    GAATTCATCA TTTTTCANGT CCTCAAGTGG ATGTTTCTCA TTTNCCATGA TTTTAAGTTT      60
65    TCTCGCCATA TTCCTGGTCC TACAGTGTGC ATTTCTCCAT TTTNCACGTT TTNCAGTGAT     120
66    TTCGTCATTT TCAAGTCCTC AAGTGGATGT TTCTCATTTN CCATGAATTT CAGTTTTCTN     180
67    GCCATATTCC ACGTCCTACA GNGGACATTT CTAAATTTNC CACCTTTTTC AGTTTTCCTC     240
68    GCCATATTTC ACGTCCTAAA ATGTGTATTT CTCGTTTNCC GTGATTTTCA GTTTTCTCGC     300
69    CAGATTCCAG GTCCTATAAT GTGCATTTCT CATTTNNCAC GTTTTTCAGT GATTTCGTCA     360
70    TTTTTTCAAG TCGGCAAGTG GATGTTTCTC ATTTNCCATG ATTTNCAGTT TTCTTGNAAT     420
71    ATTCCATGTC CTACAATGAT CATTTTTAAT TTTCCACCTT TTCATTTTTC CACGCCATAT     480
72    TTCATGTCCT AAAGTGTATA TTTCTCCTTT TCCGCGATTT TCAGTTTTCT CGCCATATTC     540
73    CAGGTCCTAC AGTGTGCATT CCTCATTTTT CACCTTTTTC ACTGATTTCG TCATTTTTCA     600
74    AGTCGTCAAC TGGATCTTTC TAATTTTCCA TGATTTTCAG TTATCTTGTC ATATTCCATG     660
75    TCCTACAGTG GACATTTCTA AATTTTCCAA CTTTTTCAAT TTTTCTCGAC ATATTTGACG     720
76    TGCTAAAGTG TGTATTCTT ATTTTCCGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGTC     780
77    CTAATAGTGT GCATTTCTCA TTTTTCACGT TTTTCAGTGA TTTCGTCATT TTTTCCAGTT     840
78    GTCAAGGGGA TGTTTCTCAT TTTCCATGAG TGTCAGTTTT CTTGCTATAT CCATGTCCT     900
79    ACAGTGACAT TTCTAAATAT TATACCTTTT TCAGTTTTTC TCACCATATT TCACGTCCTA     960
80    AAGTATATAT TTCTCATTTT CCCTGATTTT CAGTTTCCTT GCCATATTCC AGGTCCTACA    1020
81    GTGTGCATTT CTCATTTTTC ACGTTTTTCA GTAATTTCTT CATTTTTTAA GCCCTCAAAT    1080
82    GGATGTTTCT CATTTTCCAT GATTTTCAGT TTTCTTGCCA TATACCATGT CCTACAGTGG    1140
83    ACATTTCTAA ATTATCCACC TTTTTCAGTT TTTCATCGGC ACATTTCACG TCCTAAAGTG    1200
84    TGTATTTCTA ATTTTCAGTG ATTTTCAGTT TTCTCGCCAT ATTCCAGGAC CTACAGTGTG    1260
85    CATTTCTCAT TTTTCACGTT TTTCAGTGAA TTC                                 1293
86
87            (2) INFORMATION FOR SEQ ID NO:2:
88
89        (i) SEQUENCE CHARACTERISTICS:
90          (A) LENGTH: 1044 base pairs
91          (B) TYPE: nucleic acid
92          (C) STRANDEDNESS: single
93          (D) TOPOLOGY: linear
94
95        (ii) MOLECULE TYPE: Genomic DNA
96        (iii) HYPOTHETICAL: NO
97        (iv) ANTI-SENSE: NO
98        (v) FRAGMENT TYPE:
99        (vi) ORIGINAL SOURCE:
```

INPUT SET: S26847.raw

```
100            (ix)  FEATURE:
101
102            (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:2:
103
104    AGGCCTATGG TGAAAAAGGA AATATCTTCC CCTGAAAACT AGACAGAAGG ATTCTCAGAA     60
105    TCTTATTTGT GATGTGCGCC CCTCAACTAA CAGTGTTGAA GCTTTCTTTT GATAGAGCAG    120
106    TTTTGAAACA CTCTTTTTGT AAAATCTGCA AGAGGATATT TGGATAGCTT TGAGGATTTC    180
107    CGTTGGAAAC GGGATTGTCT TCATATAAAC CCTAGACAGA AGCATTCTCA GAAGCTTCAT    240
108    TGGGATGTTT CAGTTGAAGT CACAGTGTTG AACAGTCCCC TTTCATAGAG CAGGTTTGAA    300
109    ACACTCTTTT TTGTAGTATC TGGAAGTGGA CATTTGGAGC GATCTCAGGA CTGCGGTGAA    360
110    AAAGGAAATA TCTTCCAATA AAAGCTAGAT AGAGGCAATG TCAGAAACCT TTTTCATGAT    420
111    GTATCTACTC AGCTAACAGA GTTGAACCTT CCTTTGAGAG AGCAGTTTTG AAACACTCTT    480
112    TTTGTGGAAT CTGCAAGTGG ATATTTGTCT AGCTTTGAGG ATTTCGTTGG GAAACGGGAT    540
113    TACATATAAA AAGCAGACAG CAGCATTCCC AGAAACTTCT TTGTGATGTT TGCATTCAAG    600
114    TCACAGAGTT GAACATTCCC TTTCATAGAG CAGGTTTGAA ACACACTTTT TGATGTATCT    660
115    GGATGTGGAC ATTTGCAGCG CTTTCAGGCC TAAGGTGAAA AGGAAATATC TTCCCCTGAA    720
116    AACTAGACAG AAGCATTCTC AGAAACTTAT TTGTGATGTG CGCCCTCAAC TAACAGTGTT    780
117    GAAGCTTTCT TTTGATAGAG GCAGTTTTGA AACACTCTTT TGTGGAATCT GCAAGTGGAT    840
118    ATTTGTCTAG CTTTGAGGAT TTCTTTGGAA ACGGGATTAC ATATAAAAAG CAGACAGCAG    900
119    CATTCCCAGA ATCTTGTTTG TGATGTTTGC ATTCAAGTCA CAGAGTTGAA CATTCCCTTT    960
120    CAGAGAGCAG GTTTGAACAC TCTTTTTATA GTATCTGGAT GTGGACATTT GGAGCGCTTT   1020
121    CAGGGGGGAT CCTCTAGAAT TCCT                                         1044
122
123
124
125            (2)  INFORMATION FOR SEQ ID NO:3:
126
127            (i)  SEQUENCE CHARACTERISTICS:
128               (A)  LENGTH: 2492 base pairs
129               (B)  TYPE: nucleic acid
130               (C)  STRANDEDNESS: single
131               (D)  TOPOLOGY: linear
132
133            (ii)  MOLECULE TYPE: Genomic DNA
134            (iii) HYPOTHETICAL: NO
135            (iv)  ANTI-SENSE: NO
136            (v)   FRAGMENT TYPE:
137            (vi)  ORIGINAL SOURCE:
138            (ix)  FEATURE:
139
140            (xi)  SEQUENCE DESCRIPTION: SEQ ID NO:3:
141
142    CTGCAGCTGG GGGTCTCCAA TCAGGCAGGG GCCCCTTACT ACTCAGATGG GGTGGCCGAG     60
143    TAGGGAAGG  GGGTGCAGGC TGCATGAGTG GACACAGCTG TAGGACTACC TGGGGGCTGT    120
144    GGATCTATGG GGGTGGGGAG AAGCCCAGTG ACAGTGCCTA GAAGAGACAA GGTGGCCTGA    180
145    GAGGGTCTGA GGAACATAGA GCTGGCCATG TTGGGGCCAG GTCTCAAGCA GGAAGTGAGG    240
146    AATGGGACAG GCTTGAGGAT ACTCTACTCA GTAGCCAGGA TAGCAAGGAG GGCTTGGGGT    300
147    TGCTATCCTG GGGTTCAACC CCCCAGGTTG AAGGCCCTGG GGGAGATGGT CCCAGGACAT    360
148    ATTACAATGG ACACAGGAGG TTGGGACACC TGGAGTCACC AAACAAAACC ATGCCAAGAG    420
149    AGACCATGAG TAGGGGTGTC CAGTCCAGCC CTCTGACTGA GCTGCATTGT TCAAATCCAA    480
150    AGGGCCCCTG CTGCCACCTA GTGGCTGATG GCATCCACAT GACCCTGGGC CACACGCGTT    540
151    TAGGGTCTCT GTGAAGACCA AGATCCTTGT TACATTGAAC GACTCCTAAA TGAGCAGAGA    600
152    TTTCCACCTA TTCGAAACAA TCACATAAAA TCCATCCTGG AAAAAGCCTG GGGGATGGCA    660
```

INPUT SET: S26847.raw

```
153    CTAAGGCTAG GGATAGGGTG GGATGAAGAT TATAGTTACA GTAAGGGGTT TAGGGTTAGG    720
154    GATCAACGTT GGTTAGGAGT TAGGGATACA GTAGGGTACC GGTAGGGTTA GGGGTTAGGG    780
155    TTAGGGGTTA GGGTTAGGGT TAGGGTTAGG GTTAGGGTTA GGGGTTAGGG GTTAGGGTTA    840
156    GGGTTAGGTT TTGGGGTGGC GTATTTTGGT CTTATACGCT GTGTTCCACT GGCAATGAAA    900
157    AGAGTTCTTG TTTTTCCTTC AGCAATTTGT CATTTTTAAA AGAGTTTAGC AATTCTAACA    960
158    GATATAGACC AGCTGTGCTA TCTCATTGTG GTTTTCAATT GTAACCACAT TGTGGTTTCA   1020
159    ATGTGTTTAC TTGCCATCTG TAGATCTTCT TTGCGTGAGG TGTCTGTTCA GATGTGTGTG   1080
160    CATTTCTTGN NTTTNGGCTG TTTAACTTAT TGTTTAGTTT TAATAATTTT TTATATATTT   1140
161    GAAGACAAAT CTTTCTCAGA TGTGTATTTG CAAATATTTC TTCAATATGA GGCTTGCTTT   1200
162    TGTCTCTAAC AAGGTCTCTT CAGAGATAAC TTAAATATAA GAAATCCACA CTGTCACTTC   1260
163    TTTTGTGTAT ATCTACCTTT TGTGTCATTT GTTAAAATTC ATTACCAAAC CCAAAGGCAG   1320
164    ATAGCTTTTC TTCTATTGTT TCTTCTAGAA ATTTGTATAG TTTTGCATTT TTAGTGTAAG   1380
165    GATGATTTTG AGTGATTATT TGTGTAAGTT GTAAAGTTTT CGTCTATATC CATATCATTT   1440
166    CTTATGGTTT CCAATTAATC GTTCCCTCAC TATTTTTGGG AAAGACACAG GATAGTGGGC   1500
167    TTTGTTAGAG TAGATAGGTA GCTAGACATG AACAGGAGGG GGCCTCCTGG AAAAGGGAAA   1560
168    GTCTGGGAAG GCTCACCTGG AGGACCACCA AAAATTCACA TATTAGTAGC ATCTCTAGTG   1620
169    CTGGAGTGGA TGGGCACTTG TCAATTGTGG GTAGGAGGGA AAAGAGGTCC TATGCAGAAA   1680
170    GAAACTCCCT AGAACTCCTC TGAAGATGCC CCAATCATTC ACTCTGCAAT AAAAATGTCA   1740
171    GAATATTGCT AGCTACATGC TGATAAGGNN AAAGGGGACA TTCTTAAGTG AAACCTGGCA   1800
172    CCATAAGTAC. AGATTAGGGC AGAGAAGGAC ATTCAAAAGA GGCAGGCGCA GTAGGTACAA   1860
173    ACGTGATCGC TGTCAGTGTG CCTGGGATGG CGGGAAGGAG GCTGGTGCCA GAGTGGATTC   1920
174    GTATTGATCA CCACACATAT ACCTCAACCA ACAGTGAGGA GGTCCCACAA GCCTAAGTGG   1980
175    GGCAAGTTGG GGAGCTAAGG CAGTAGCAGG AAAACCAGAC AAAGAAAACA GGTGGAGACT   2040
176    TGAGACAGAG GCAGGAATGT GAAGAAATCC AAAATAAAAT TCCCTGCACA GGACTCTTAG   2100
177    GCTGTTTAAT GCATCGCTCA GTCCCACTCC TCCCTATTTT TCTACAATAA ACTCTTTACA   2160
178    CTGTGTTTCT TTTCAATGAA GTTATCTGCC ATCTTTGTAT TGCCTCTTGG TGAAAATGTT   2220
179    TCTTCCAAGT TAAACAAGAA CTGGGACATC AGCTCTCCCC AGTAATAGCT CCGTTTCAGT   2280
180    TTGAATTTAC AGAACTGATG GGCTTAATAA CTGGCGCTCT GACTTTAGTG GTGCAGGAGG   2340
181    CCGTCACACC GGGACCAAGA GTGCCCTGCC TAGTCCCCAT CTGCCCGCAG GTGGCGGCTG   2400
182    CCTCGACACT GACAGCAATA GGGTCCGGCA GTGTCCCCAG CTGCCAGCAG GGGGCGTACG   2460
183    ACGACTACAC TGTGAGCAAG AGGGCCCTGC AG                                 2492
184
185              (2) INFORMATION FOR SEQ ID NO:4:
186
187          (i) SEQUENCE CHARACTERISTICS:
188             (A) LENGTH: 28 base pairs
189             (B) TYPE: nucleic acid
190             (C) STRANDEDNESS: single
191             (D) TOPOLOGY: linear
192
193          (ii) MOLECULE TYPE: Genomic DNA
194          (iii) HYPOTHETICAL: NO
195          (iv) ANTI-SENSE: NO
196          (v) FRAGMENT TYPE:
197          (vi) ORIGINAL SOURCE:
198          (ix) FEATURE:
199
200          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:
201
202    GGGGAATTCA TTGGGATGTT TCAGTTGA                                      28
203
204              (2) INFORMATION FOR SEQ ID NO:5:
205
```

```
206          (i) SEQUENCE CHARACTERISTICS:
207            (A) LENGTH: 29 base pairs
208            (B) TYPE: nucleic acid
209            (C) STRANDEDNESS: single
210            (D) TOPOLOGY: linear
211
212          (ii) MOLECULE TYPE: Genomic DNA
213          (iii) HYPOTHETICAL: NO
214          (iv) ANTI-SENSE: NO
215          (v) FRAGMENT TYPE:
216          (vi) ORIGINAL SOURCE:
217          (ix) FEATURE:
218
219          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:
220
221   CGAAAGTCCC CCCTAGGAGA TCTTAAGGA                          29
222
223          (2) INFORMATION FOR SEQ ID NO:6:
224
225          (i) SEQUENCE CHARACTERISTICS:
226            (A) LENGTH: 47 base pairs
227            (B) TYPE: nucleic acid
228            (C) STRANDEDNESS: single
229            (D) TOPOLOGY: linear
230
231          (ii) MOLECULE TYPE: RNA
232          (iii) HYPOTHETICAL: NO
233          (iv) ANTI-SENSE: NO
234          (v) FRAGMENT TYPE:
235          (vi) ORIGINAL SOURCE:
236          (ix) FEATURE:
237
238          (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:
239
240   CCGCTTAATA CTCTGATGAG TCCGTGAGGA CGAAACGCTC TCGCACC        47
241
242
243
244          (2) INFORMATION FOR SEQ ID NO:7:
245
246          (i) SEQUENCE CHARACTERISTICS:
247            (A) LENGTH: 25 base pairs
248            (B) TYPE: nucleic acid
249            (C) STRANDEDNESS: single
250            (D) TOPOLOGY: linear
251
252          (ii) MOLECULE TYPE: Genomic DNA
253          (iii) HYPOTHETICAL: NO
254          (iv) ANTI-SENSE: NO
255          (v) FRAGMENT TYPE:
256          (vi) ORIGINAL SOURCE:
257          (ix) FEATURE:
258
```

SEQUENCE VERIFICATION REPORT
PATENT APPLICATION   *US/09/096,648*

INPUT SET: S26847.raw

Line	Error	Original Text